(19) Europäisches Patentamt
Eur pean Pat nt Office
Office européen des brevets

(11) EP 0 786 521 A1

# (12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
**30.07.1997 Bulletin 1997/31**

(21) Application number: **96119668.0**

(22) Date of filing: **12.02.1992**

(51) Int. Cl.[6]: **C12N 15/31**, C12N 15/74, C12N 15/70, C07K 14/315, C12N 1/21, A61K 39/09, C12Q 1/68

(84) Designated Contracting States:
**AT BE CH DE DK ES FR GB GR IT LI LU MC NL SE**

(30) Priority: **15.02.1991 US 656773**

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
**92907062.1 / 0 571 538**

(71) Applicant: **UAB RESEARCH FOUNDATION**
**Birmingham, AL 35294-009 (US)**

(72) Inventors:
- **Briles, David E.**
  **Birmingham, Alabama 35222 (US)**
- **Yother, Janet L.**
  **Birmingham, Alabama 35242 (US)**
- **McDaniel, Larry S.**
  **Birmingham, Alabama 35210 (US)**

(74) Representative: **Smart, Peter John**
**W.H. BECK, GREENER & CO**
**7 Stone Buildings**
**Lincoln's Inn**
**London WC2A 3SZ (GB)**

Remarks:
This application was filed on 07 - 12 - 1996 as a divisional application to the application mentioned under INID code 62.

## (54) Nucleic acid encoding PspA

(57) Nucleic acid encoding pneumococcal surface protein A is disclosed both as the full sequence and as primers and probes derived therefrom, together with vectors containing the nucleic acid and methods of obtaining PspA by expression of said nucleic acid and PspA so produced.

EP 0 786 521 A1

Printed by Rank Xerox (UK) Business Services
2.14.11/3.4

## Description

### FIELD OF INVENTION

The present invention is concerned with the development of an improved vaccine against pneumococcal infections.

### BACKGROUND TO THE INVENTION

Streptococcus pneumoniae is an important cause of otitis media, meningitis, bacteremia and pneumonia. Despite the use of antibiotics and vaccines, the prevalence of pneumococcal infections has declined little over the last twenty-five years.

It is generally accepted that immunity to Streptococcus pneumoniae can be mediated by specific antibodies against the polysaccharide capsule of the pneumococcus. However, neonates and young children fail to make an immune response against polysaccharide antigens and can have repeated infections involving the same capsular serotype.

One approach to immunising infants against a number of encapsulated bacteria is to conjugate the capsular polysaccharide antigens to protein to make them immunogenic. This approach has been successful, for example, with Haemophilus influenzae b (see U.S. Patent No. 4,496,538 to Gordon and U.S. Patent No. 4,673,574 to Anderson). However, there are over eighty known capsular serotypes of S. pneumoniae of which twenty-three account for most of the disease. For a pneumococcal polysaccharide-protein conjugate to be successful, the capsular types responsible for most pneumococcal infections would have to be made adequately immunogenic. This approach may be difficult, because the twenty-three polysaccharides included in the presently-available vaccine are not all adequately immunogenic, even in adults.

An alternative approach for protecting children, and also the elderly, from pneumococcal infection would be to identify protein antigens that could elicit protective immune responses. Such proteins may serve as a vaccine by themselves, may be used in conjunction with successful polysaccharide-protein conjugates, or as carriers for polysaccharides.

In McDaniel et al (I), J.Exp.Med. 160:386-397, 1984, there is described the production of hybridoma antibodies that recognize cell surface polypeptide(s) on S. pneumoniae and protection of mice from infection with certain strains of encapsulated pneumococci by such antibodies. This surface protein antigen has been termed "pneumococcal surface protein A" or PspA for short.

In McDaniel et al (II), Microbial Pathogenesis 1:519-531, 1986, there are described studies on the characterization of the PspA. Considerable diversity in the PspA molecule in different strains was found, as were differences in the epitopes recognized by different antibodies.

In McDaniel et al (III), J.Exp.Med. 165:381-394, 1987, there is disclosed that immunization of X-linked immunodeficient (XID) mice with non-encapsulated pneumococci expressing PspA, but not isogenic pneumococci lacking PspA, protects mice from subsequent fatal infection with pneumococci.

In McDaniel et al (IV), Infect. Immun., 59:222-228, 1991, there is described immunization of mice with a recombinant full length fragment of PspA that is able to elicit protection against pneumococcal strains of capsular types 6A and 3.

In Crain et al, Infect.Immun., 56:3293-3299, 1990, there is described a rabbit antiserum that detects PspA in 100% (n = 95) of clinical and laboratory isolates of strains of S. pneumoniae. When reacted with seven monoclonal antibodies to PspA, fifty-seven S. pneumoniae isolates exhibited thirty-one different patterns of reactivity.

The PspA protein type is independent of capsular type. It would seem that genetic mutation or exchange in the environment has allowed for the development of a large pool of strains which are highly diverse with respect to capsule, PspA, and possibly other molecules with variable structures. Variability of PspA's from different strains also is evident in their molecular weights, which range from 67 to 99 kD. The observed differences are stably inherited and are not the result of protein degradation.

Immunization with a partially purified PspA from a recombinant λ gt11 clone, elicited protection against challenge with several S. pneumoniae strains representing different capsular and PspA types, as described in McDaniel et al (IV), Infect. Immun. 59:222-228, 1991. Although clones expressing PspA were constructed according to that paper, the product was insoluble and isolation from cell fragments following lysis was not possible.

While the protein is variable in structure between different pneumococcal strains, numerous cross-reactions exist between all PspA's, suggesting that sufficient common epitopes may be present to allow a single PspA or at least a small number of PspA's to elicit protection against a large number of S. pneumoniae strains.

In addition to the published literature specifically referred to above, the inventors, in conjunction with co-workers, have published further details concerning PspA's, as follows:

1. Abstracts of 89th Annual Meeting of the American Society for Microbiology, p. 125, item D-257, May 1989;
2. Abstracts of 90th Annual Meeting of the American Society for Microbiology, p. 98, item D-106, May 1990;

3 Abstracts of 3rd International ASM Conference on Streptococcal Genetics, p. 11, item 12, June 1990;
4. Talkington et al, Infect. Immun. 59:1285-1289, 1991;
5. Yother et al, J. Bacteriol. 174:601-609, 1992; and
6. Yother et al, J. Bacteriol. 174:610-618, 1992.

The latter three publications occurred after the filing of the aforesaid USSN 656,773.

In the specification which follows and the drawings accompanying the same, there are utilized certain accepted abbreviations with respect to the amino acids represented thereby. The following Table I identifies those abbreviations:

TABLE I

| AMINO ACID ABBREVIATIONS | |
|---|---|
| A = Ala = Alanine | M = Met = Methionine |
| C = Cys = Cysteine | N = Asn = Asparagine |
| D = Asp = Aspartic Acid | P = Pro = Proline |
| E = Glu = Glutamic Acid | Q = Gln = Glutamine |
| F = Phe = Phenylalanine | R = Arg = Arginine |
| G = Gly = Glycine | S = Ser = Serine |
| H = His = Histidine | T = Thr = Threonine |
| I = Ile = Isoleucine | V = Val = Valine |
| K = Lys = Lysine | W = Try = Tryptophan |
| L = Leu = Leucine | Y = Tyr = Tyrosine |

## SUMMARY OF INVENTION

The present invention relates to the preparation of mutants of S. pneumoniae that secrete an immunogenic truncated form of the PspA protein, and isolation and purification of the secreted protein. The truncated form of the PspA protein is immunoprotective and contains the protective epitopes of PspA. The PspA protein is soluble in physiologic solution and lacking at least the functional cell membrane anchor region.

## BRIEF DESCRIPTION OF DRAWINGS

Figure 1 is a schematic representation of the domains of the mature PspA;
Figure 2 is the N-terminal amino acid sequence of PspA, wherein bold upper case letters denote charged hydrophilic amino acids, lower-case letters designate apolar, hydrophobic residues, and underlined bold lower case letters denote uncharged, polar, hydrophilic residues;
Figure 3 is the DNA sequence of the pspA gene with deduced amino acid sequence for the PspA protein;
Figure 4 depicts the restriction map of pspA (Figure 4A) and the use of insertion-duplication mutagenesis to construct mutations in the pspA gene (Figure 4B),
Figure 5 shows the deduced amino acid sequence for the N-terminal region of PspA and the general location of epitopes recognized by monoclonal antibodies;
Figure 6 shows antibody reactivity with PspA fragments produced by various pspA gene segments; and
Figure 7 shows the mapped location of epitopes in the PspA fragments produced by the different pspA gene segments.

## GENERAL DESCRIPTION OF INVENTION

According to one aspect of the present invention, there is provided a purified immunoprotective pneumococcal surface protein, comprising a truncated form of PspA which contains the immunoprotective epitopes of the protein and up to about 90% of the whole PspA protein and from which the functional cell membrane anchor region is absent.

Through the technique of insertion-duplication mutagenesis of the pspA gene of the strain Rx1 of Streptococcus pneumoniae with plasmids containing cloned fragments of the pspA structural gene, it has been possible to produce

soluble fragments of PspA that are secreted by pneumococci.

In another aspect of the present invention, therefore, there is provided a method of forming an immunoprotective truncated PspA protein, which comprises effecting insertion-duplication mutagenesis of a bacterium with a pspA gene resulting in the coding of a truncated expressible PspA protein, growing the mutated bacterium to effect expression of a truncated PspA protein, and isolating the protein.

The molecular size of the purified truncated PspA protein obtained may be varied by directing the point of insertion, which determines the termination of gene expression, to different points in the pspA gene. For example, an N-terminal fragment of apparent molecular weight of 43 kD, constituting approximately one-half of the native protein, has been found useful.

The truncated segment which is produced by this procedure is capable of eliciting protection in mice from fatal challenge with type 3 S. pneumoniae, demonstrating for the first time that a purified PspA can elicit protection and that this truncated segment of the protein contains protective epitopes of PspA.

Amino acid sequence information was obtained on the N-terminal 45 amino acids of the truncated segment of PspA. This sequence is shown in Figure 2. Predictive secondary structural analysis shows that this sequence has a very strong alpha-helical formation, with no non-helical inserts. About 51% of the segment is composed only of two amino acids, namely lysine, a charged amino acid, and alanine, a non-polar amino acid.

Analysis of this 45-amino acid sequence also reveals that it contains a seven-residue periodicity (see Figure 2). In PspA, the periodicity begins with residue 8 and extends throughout the entire sequence, for nearly eleven turns of the helix. Positions "a" and "d" are occupied by apolar amino acids and position "b", "c" and "f" generally contain hydrophilic amino acids. Position "f" is predominantly occupied by lysine. Having regard to these observations, this region of PspA is very likely in an alpha-helical coiled-coil configuration. The deduced amino acid sequence for the whole of the α-helical coiled-coil region is shown in Figure 5.

We also have cloned and sequenced the entire coding region of pspA (see Figure 3). The deduced amino acid sequence for the PspA protein reveals three distinct regions of the PspA molecule, shown schematically in Figure 1. Accordingly, a further aspect of the present invention, there is provided a biologically-pure recombinant DNA molecule coding for the PspA protein or portions thereof and having a coding sequence included within set forth in Figure 3 or having substantial homology thereto.

The DNA sequence of the pspA gene is contained on a HindIII - KpnI fragment that is 2086 base pairs in length. The pspA gene itself represents approximately 1985 base pairs of this fragment, and comprises an initial region containing transcription and translational signals with translation starting at the ATG/met (nucleotide position 127, codon position -31), followed by a leader sequence extending from the AAG/met (nucleotide position 127, codon position -31) to CGA/ala (nucleotide position 217, codon -1). Mature Pspa starts with the glu amino acid at nucleotide position 220 (codon +1) and ends at the translational stop TAA/OCH at nucleotide position 1984. This translational stop codon is followed by transcription termination signals.

The amino terminal of the protein sequence, predicted from the DNA sequence of Figure 3, contains a 31 amino acid leader sequence and a 45 amino acid sequence identical to the 45 amino acid sequence of the N-terminal of PspA (Figure 2). The amino end of the predicted protein sequence is highly charged and α-helical in nature. This region has homology with tropomyosin at the amino acid level (approximately 22% identity and 50% similarity). This homology is due largely to a repeating seven residue periodicity where the first and fourth amino acids are hydrophobic, the intervening amino acids are helix-promoting and the seventh amino acid is charged. This pattern is consistent with that of an α-helical coiled-coil molecule and indicates that the α-helical coil extends through the N-terminal half of the molecule. The amino acid sequence of the whole of the α-helical coil region is shown in Figure 5.

Following the charged helical region is a proline-rich region in which 23 of 81 amino acids are prolines. Immediately carboxy to the proline-rich region is the first of ten highly homologous twenty amino acid repeats. The only significantly hydrophobic region in the sequenced portion of the molecule begins at the last repeat. This potential membrane-spanning region contains several charged amino acids preceding the translational stop codon.

The insertionally-inactivated mutants of S.pneumoniae lacking the C-terminal anchor regions are capable of growth in chemically-defined medium and secrete the N-terminal portion of the PspA protein into the medium. The N-terminal region of PspA is highly soluble in the culture medium and is much easier to isolate than the entire molecule. Soluble truncated molecules have been produced using insertional duplicational mutagenesis directed by the cloned PspA DNA fragments shown in Figure 4. Expression of the same truncated construct (with the pneumococcal promoter) in E.coli results in the same PspA fragment being secreted into the periplasm of E.coli. PspA is readily released from the periplasm by hypotonic lysis.

Truncated PspA is isolated from culture medium of mutant pneumococci in any convenient manner, such as by tangential flow filtration. Ion-exchange chromatography then is performed on an anionic resin to purify the protein. In this procedure, the solution containing PspA is dialyzed to pH6 in 0.02 M salt solution and passed over the resin. The PspA is eluted from the resin with a gradient of 0.08 to 2.0 M ionic strength and is collected in the fraction between 0.34 and 0.87 M ionic strength, depending on the nature of the column used.

The PspA may be further purified by sodium dodecyl sulfate polyacrylamide gel (SDS-PAGE) electrophoresis. The

PspA-containing portion of the gel is identified by staining of the gel and PspA is electroeluted from this portion.

The electrophoresis purification is convenient when only small quantities of PspA are being handled. As an alternative, more suited to large-scale production, the protein may be purified by size chromatography in a pH7 phosphate buffer.

Since it is possible to obtain expression of the truncated form of PspA into the culture medium, as opposed to it being trapped within the cell wall and making purification much more complicated, it is possible to isolate other proteins that have been cloned into the truncated pspA gene by making fusion proteins between PspA and other proteins. Such a technique may be employed to enhance the immunogenicity or preserve the immunogenic structural conformation or presentation of the gene product, to permit the fusion protein to be used in immunization, which may be systemic and/or mucosal, against disease.

One example of such a fusion protein is a fusion of the soluble N-terminal region of PspA and the B-subunit of cholera toxin. Fusion proteins also may be formed by chemical attachment of the truncated PspA protein to other proteins.

Another aspect of the present invention, therefore, provides a method for the production of cloned proteins, which comprises fusing a pspA gene coding for a truncated form of PspA protein with the gene coding for another protein to form a fusion protein clone, transforming S.pneumoniae, E.coli or other bacteria with the fusion protein clone, growing the transformed bacterium to effect expression of a fusion protein comprising truncated PspA and the other protein into the culture medium, and isolating the fusion protein.

By using this technique, there can be produced cloned proteins in gram positive bacteria, such as pneumococci, for example, S.pneumoniae, and mycobacteria, for example, Bacille Calmette-Guerin (BCG). This approach overcomes the problems inherent in the production of proteins in gram negative bacteria, such as E. coli, usually used for cloning, in particular the need to purify the recombinant proteins from endotoxin and the toxicity of many gram positive DNA sequences in gram negative organisms.

For the expression of a fusion protein comprising the soluble N-terminal region of PspA and the B-subunit of cholera toxin (CTB), a gene fusion of a pspA gene coding for a truncated form of PspA protein with a ctxB gene coding for the B-subunit of cholera toxin is effected. Following expression of the fusion protein, the PspA and CTB may be cleaved one from another by dilute acid at an asparagine-proline sequence, known to be labile to dilute acid, engineered at the fusion site of the two proteins.

CTB is known to be highly specific for monosinloganglioside ($G_{M1}$). Accordingly, the fusion PspA-CTB protein may be isolated from the culture medium by adsorption to a $G_{M1}$ affinity column, from which the fusion protein subsequently may be eluted at low pH.

The PspA-CTB fusion protein finds considerable utility in solid phase immunoadsorbant assays. By using the fusion protein, it is possible to coat solid supports, such as microtitration plates, with PspA fragments without having first to isolate the PspA fragments. This may be done by adding bacterial extract containing the fusion protein to plates coated with $G_{M1}$. The PspA-CTB fusion protein then binds to $G_{M1}$ through the CTB moiety, thereby coating the solid support with PspA. The resulting coated product then may be used in a solid phase immunoadsorbant assay for the detection of PspA antibody and/or antigen in test samples. Such immunoadsorbant assays constitute an additional aspect of this invention.

The PspA attachment/anchor region, containing the proline-rich region, the repeat region and/or the C-terminus of PspA, also may be employed to effect expression of heterologous proteins in pneumococci, or other gram positive or gram negative bacteria in the which the attachment/anchor region is functional. Generally, expression is effected on bacterial membrane, cell walls or cell surfaces in gram positive bacteria and in the periplasm of gram negative bacteria. An example of such heterologous protein is the B-subunit of cholera toxin.

As mentioned above, the truncated form of PspA provided herein contains the immunoprotective epitopes of the protein and hence is useful in a vaccine against pneumococcal infection. Accordingly, a yet further aspect of the present invention provides a vaccine against pneumococcal infection comprising, as an immunogenically-active component, the purified immunoprotective pneumococcal surface protein provided herein. The PspA protein may be employed as one component of a multicomponent vaccine which is effective in providing protection from a variety of infections.

In addition, gram positive bacteria which have been transformed to express the pspA gene coding for the truncated soluble PspA protein may be employed, in a live-attenuated or killed form, as an immunologically-active component of a vaccine against pneumococcal infection. In the transformed bacterium, such pspA gene may be fused to a gene coding for another protein. Accordingly, an additional aspect of this invention provides a vaccine against pneumococcal infection comprising, as an immunologically-active component, a live-attenuated or killed bacteria containing a gene coding for the truncated form of PspA.

The truncated form of PspA also may be employed in conjugates with normally weakly-immunogenic or non-immunogenic protection-eliciting molecules, such as various polysaccharides, to achieve immunogenic potentiation thereof. An additional aspect of the invention, therefore, provides a vaccine comprising, as an immunogenically-active component, a conjugate of the purified immunoprotective pneumococcal surface protein provided herein and a normally weakly-immunogenic or non-immunogenic protection-eliciting molecule.

Conserved sequences of pspA, particularly those in the proline-rich and/or repeat regions of the gene, may be

employed as probes to detect the presence of pneumococci of various strains, through detection of pneumococcal DNA, in tissues, body fluids and/or secretions. Similarly, portions of the pspA gene may be used in diagnostic kits for the detection of pneumococcal infections.

In addition, primers made based on conserved sequences of pspA, particularly those in the proline-rich and/or repeat regions, may be used to assay for the presence of pneumococci in tissues, body fluids and/or secretions, through amplification of pneumococcal DNA. In this regard, a single primer pair derived from the nucleotide sequence of the pspA gene of S.pneumoniae may be employed in an assay using the polymerase chain reaction (PCR) for the specific detection of Streptococcus pneumoniae.

Specific amplification has been achieved of a 678 base pair DNA fragment from S.pneumoniae strain Rx1. After 30 cycles of amplification, the amplimer was detectable by agarose gel electrophoresis. The fragment was successfully amplified in all 32 strains of S.pneumoniae tested. PCR DNA amplification was able to detect less than an estimated 20 ficograms total genomic pneumococcal DNA.

Primers LSM1 and LSM2, having the nucleotide sequences:
LSM1 5'-CCGGATCCAGCTCCTGCACCAAAAC-3'
LSM2 5'-GCGCTGCGACGGCTTAAACCCATTCACCATTGG-3'
amplified the 678 base pair product from pspA from nucleotides 1312 to 1990 of the Rx1 pspA sequence (Figure 3).

The PCR analysis using the primers described herein is performed in accordance with conventional PCR techniques, such as are described in the literature, for example, as described in Arnhem et al at C&EN Special Report, 36, October 1, 1990. For detection purposes, the primer may be labelled or labelled nucleotide triphosphates may be included in the PCR reaction to label the PCR amplification product.

The PCR primers may be prepared by well-known methods, for example, by oligonucleotide synthesis or by fragmentation of a larger nucleotide sequence using suitable restriction enzymes.

The ability to use a single primer capable of detecting a large number of S.pneumoniae strains enables a universal PCR detection kit to be provided which is able to diagnose pneumococcal infection in mammals, including humans, independent of the strain which has caused the disease.

## STRAINS, PLASMIDS AND PROBES

In the Examples which follow as well as in the accompanying drawings, reference is made to certain plasmids and bacterial strains transformed by such plasmids as well as vector DNA segments, some of which have been deposited with ATCC and all of which are fully described herein. The following Table II provides a summary of such materials.

TABLE II

| Identification | Type | Description | Deposit | Location |
|---|---|---|---|---|
| JY4313 | E.coli strain | PspA DNA | ATCC 68529 | Fig. 1 |
| JY2008 | S.pneumomiae strain | PspA fragment 43 kDa | ATCC 55143 | Fig. 1 |
| JY4306 | E.coli strain | PspA fragment 43 kDa | ATCC 68522 | Fig. 3 |
| JY4310 | | PspA fragment 21 kDa | None | Fig. 7 |
| JY4285 | | PspA fragment 18 kDa | None | Fig. 7 |
| pJY4163 | Plasmid | Expression plasmid used for expression of PspA -CTB fusion protein (29 kDa) | None | Fig. 6 |
| JY4323 | DNA probe | HindIII-KpaI segment | None | Fig. 9 |
| JY4306 | DNA probe | HindIII-Dra-I segment | None | Fig. 9 |
| JY4262 | DNA probe | BclI-Bst-NI segment | None | Fig. 9 |

## EXAMPLES

### Example 1

This Example illustrates the preparation and growth of novel strains of S. pneumoniae.

The S. pneumoniae strain Rx1, which is a non-encapsulated derivative of capsular type 2 strain D39 (National Col-

lection of Type Cultures, London, NCTC #7466), was subjected to insertional inactivation (as described in McDaniel et al (III) 1987, Crain et al 1990, Talkington et al 1991, with 10 different cloned fragments of PspA (see Figure 4). These fragments have all been obtained from restriction digests of cloned PspA DNA on a plasmid in E.coli strain JY4313 (deposited with the American Type Culture Collection on January 31, 1991 under ATCC accession number 68529). This insertional duplication mutagenesis (see Figure 4) results in the termination of gene expression near the 3' end of the cloned fragment.

One of the resultant strains, JY2008 (deposited with the American Type Culture Collection on January 24, 1991 under accession number 55143), which was produced by a fragment of DNA encoded in pKSD300 (McDaniel et al (III) 1987), produces a PspA fragment of 27 kDa (apparent molecular weight 43 kDa). This fragment is approximately 40% the size of the native 65 kDa (84 kDa apparent size) protein.

The expected molecular size is based on the deduced amino acid sequence and the apparent molecular size is based on migration in SDS-PAGE. The difference between expected and apparent molecular size is due to the conformation of the PspA fragment.

The proline and repeats/anchor regions (see Figure 1) were deleted and the resulting protein was unable to attach to cell due to their absence. The unattached protein then may be isolated from culture supernatants, as described below.

By directing the insertion to different points in the pspA gene, different lengths of truncated, non-attached PspA protein derivatives can be produced, as seen in Figure 7.

The pneumococcal strain JY2008 was grown in 6 liters of a chemically defined medium (see Inf. Imm. 27:444) supplemented with 0.10% choline chloride, 0.075% L-cysteine hydrochloride and 0.25% $NaHCO_3$. The supernatant fluid of the mid-log phase culture of JY2008 was harvested using a 0.22 μm membrane tangential flow filter and concentrated 60 fold.

Introduction of the plasmid pKSD300 into the unmodified D39 strain similarly yielded the 43 kD truncated PspA protein. Introduction of the plasmid pKSD300 into the type 3 S.pneumoniae strain WU2 (PspA protein approximately 92 kD) yielded, upon growth of the organism, a non-attached truncated PspA protein of approximately 46 kD molecule size.

Example 2

This Example illustrates the purification of PspA.

The concentrated supernatant fluid, produced as described in Example 1, was washed in 0.1 M PBS, pH 7.2, and ultracentrifuged at 196,000 xg. The supernatant fluid was diluted 1:5 in 20 mM L-histidine buffer-NaCl, the pH adjusted to 6.0 and the injected into a DEAE-fibered Isonet-D2 an ion exchange column.

A stepwise NaCl gradient from 80 mM to 2 M was applied to the column and PspA-containing fractions (0.32 to 0.64 M ionic strength) were pooled and separated on an SDA-polyacrylamide gel. The proteins on a representative section of the gel were stained with Comassie Blue R-250 to identity PspA. The fraction containing PspA was excised from the remainder of the SDS-gel and electroeluted from the excised gel. The eluted protein was precipitated in a 50:50 methanol:acetone solvent and resuspended in PBS. Purity of product was confirmed by silver staining and Western Immunoblotting with mAb Xi126 (IgG 2b, k, see McDaniel et al (I), supra).

Example 3

This Example illustrates the isolation of PspA from the periplasmic space of Escherichia coli.

Isolation from the periplasmic space of E. coli was accomplished by standard techniques. E. coli strain JY4306 (which produces the 43 kDa N-terminal fragment of PspA, the amino acid sequence of which is shown in Figure 3. This strain was deposited with ATCC on January 31, 1991 under accession number 68522) was washed in buffered saline, incubated in 20% sucrose, 10mM EDTA, 25mM Tris pH 7.7 for 10 minutes at 0°C. The cells than were spun at 400 xg for 10 minutes at 0°C. All supernatant was removed from the pellet and the pellet was resuspended rapidly in about 100 volumes of 4°C water. After 10 minutes the suspension was centrifuged at 4,000 xg for 10 minutes at 4°C. The pellet was discarded and the supernatant, which contained the PspA was saved. Concentration of the supernatant was by standard procedures such as concentration against solid sucrose or ultrafiltration. Purification of the protein isolated from E. coli proceeded by the same chromatography techniques used for the isolated of the 43 kDa (truncated) PspA from the media of growing pneumococci.

Example 4

This Example illustrates the immunogenic properties of the PspA protein.

Sixteen 7-week old CBA/N mice carrying the Xid mutation (Jackson Laboratories, Bar Harber, ME) were bled via the periorbital sinus to establish pre-exposure levels of antibody to PspA. Purified PspA, prepared as described in

Example 2, was emulsified in complete Freund's adjuvant and injected subcutaneously into the inguinal and axillary regions, delivering approximately 5 μg of protein per mouse. Fourteen days later, the mice were injected intraperitoneally with 5 μg of PspA, prepared as described in Example 2. Control mice were immunized via the same routes with sterile SDS buffer. Seven days after the last immunization, all mice were bled via the periorbital sinus and were challenged intravenously with 300 CFU of the type 3 strain WU2, grown as described in Example 1.

Preimmunization and prechallenge sera were analyzed by Western immunoblots to establish baseline and postimmunization response to the truncated protein. The PspA of strain WU2 was electrophoresed and transferred to nitrocellulose membranes. The membranes were separated into strips and probed with the appropriate mouse antisera at a 1:50 dilution for 2 hours, incubated with biotinylated goat anti-mouse immunoglobulin for 1 hr, washed and incubated with Strepavidin-conjugated phosphatase. The membranes were developed with 5-bromo-4-chloro-3-indoyl phosphate toludine salt with 0.01% into blue tetrazolium.

Of the eight CBA/N mice immunized with the purified fragment of PspA, all were still alive 14 days after challenge with strain WU2 and none showed any signs of illness following challenge. Of the eight mice immunized with buffer controls, six were dead by two days post challenge, while the two remaining control mice appeared very sick, with ruffled fur, arched back and decreased movement, two to three days following challenge but survived. Chi-square analysis indicated that there was a significant difference ($P < 0.003$) in survival between the immunized and control groups.

Preimmunization and prechallenge sera were analyzed by Western immunoblotting. None of the preimmunization sera contained antibody to truncated PspA. Postimmunization sera from eight of eight mice contained detectable antibodies to PspA, and six mice had very strong anti-PspA reactions. When the challenge strain WU2 was probed with the antisera, all the immunized mice had antibodies that were highly cross-reactive with the WU2 PspA epitopes. No control mice developed antibodies to PspA.

The immunization data is summarized in the following Table III:

Table III

| Immunogen | Detection of Antibody to PspA | Alive at 2 days post challenge | Alive at 14 days post challenge |
|---|---|---|---|
| Isolated PspA (Example 2) | 8/8 | 8/8 | 8/8 |
| Sterile SDS (control) | 0/8 | 2/8 | 2/8 |

As may be seen from the data in Table III, immunization with two 5 μg doses of the purified PspA molecule elicited protection against fatal infection of CBA/N mice and elicited antibodies reactive with the PspA of the challenge strain.

Example 5

This Example illustrates sequencing of the PspA protein.

Purified PspA, prepared as described in Example 2, was electrophoresed through 9% resolving gels containing recrystallized SDS with the Laemmli buffer system (Nature 227:680). The gels were soaked twice in a 10mM 3-(cyclohexylamino)-1-propanesulfonic acid, pH 11.0, containing 10% methanol for 10 minutes. A polyvinylidene difluoride membrane (PVDF) was wetted completely for several seconds in 100% methanol, then washed in CAPS buffer for 10 min. PspA was electrotransferred to the PVDF membrane in CAPS buffer at 0.5 A for 1 hr. After transfer, the membrane was washed two times in deionized water for 5 min, and stained with 0.1% Coomassie Blue R-250 in 50% methanol for 20 minutes. The section of the membrane containing PspA was excised and destained in 40% methanol and 10% acetic acid for 5 min. The membrane was cut into small segments and stored in sterile Eppendorf tubes until sequencing.

The isolated PspA was sequenced directly from the PVDF membranes. Figure 2 depicts the N-terminal 45 residue amino acid sequence and Figure 5 depicts the amino acid sequence for the whole alpha-helical region. The DNA sequence of the whole pspA gene and the deduced amino acid sequence for the PspA protein are shown in Figure 3.

Example 6

This Example illustrates the use of the pspA 5'-sequence and/or the PspA N-terminal region to serve as an expression and leader sequence for expressing and/or excreting/secreting heterologous proteins from S.pneumoniae and E.coli. In this Example, there is described the expression of the N-terminal of the PspA protein fused to the B-subunit of cholera toxin (CTB) through a genetic fusion and the excretion of the fused protein from pneumococci and its secretion into the periplasmic space of E.coli.

A fusion protein consisting of CTB and the N-terminal half of PspA was constructed and expressed in E.coli. The HindIII/DraI pspA gene fragment used contained all the pspA transcription and translation initiation signals and the PspA signal peptide leader sequence for transport across the cell membrane. The mature PspA encoded by this fragment is predicted to be a product of 29 kDa (observed molecular weight of 42 kDa), encompassing more than 90% of the α-helical coiled-coil domain. The CTB fragment used lacked transcription and translation initiation signals. Expression from pspA promoter through pspA and then in-frame translational readthrough into the CTB-encoding gene ctxB resulted in production of a 12 kDa CTB product fused to the upstream PspA product. The PspA-CTB fusion protein was stably expressed in both high and low copy number plasmids (pUC18, more than 100 copies/cell; pJY4163, about 15 to 30 copies/cell) in E.coli.

The fusion products were of the expected size (about 54 kDa) and reacted with antibody to both PspA and CTB. That the CTB product retained its functionality was demonstrated by the ability of the fusion protein to bind ganglioside $G_{M1}$, a property of CTB.

The high level of expression of the fusion product apparently resulted in a reduced rate of processing and/or conformational changes that prevented the protein from being completely transported to the periplasm. However, in the lower copy number construct, about 60% of the fusion protein was localized in the periplasm, where large quantities were readily released from E. coli by osmotic shock.

In addition to expression in E.coli, the fusion protein also was expressed in S.pneumoniae by transformation of the low copy number construct into the avirulent S.pneumoniae Rx1 to generate an insertion-duplication mutant. In this way, the gene encoding the fusion protein was integrated into the S.pneumoniae chromosome, from which it was stably expressed. As in the case of Example 1, the truncated PspA molecule lacking the attachment/anchor region, this time in the form of the PspA-CTB fusion protein, was excreted into the culture supernatant. The fusion protein product was of the expected molecular size (54 kDa), reacted with antibody to PspA and CTB, and bound $G_{M1}$.

Example 7

This Example illustrates the use of PspA attachment or anchor region to permit expression of heterologous proteins on the surface of S.pneumoniae or other bacteria in which the attachment/anchor sequence is functional in particular the expression of a PspA-CTB (cholera toxin B subunit) fusion expressed on the surface of pneumococci.

The N-terminal encoding region of PspA, including its transcription and translation initiation signals and its signal peptide leader sequence, is linked via a translationally in-frame genetic fusion to the CTB-encoding ctxB fragment that lacks transcription and translation initiation and termination signals. This sequence is followed in-frame by the PspA attachment/anchor domain, including part or all of the proline, repeat and C-terminal domains. The resulting fusion protein is directed to the outside of the cell via the PspA leader sequence, which is cleaned following transport across the membrane, and then attached to cell by the PspA attachment/anchor sequence. The heterologous protein, located between the two PspA fragments is expressed on the outside surface of the membrane and, in S.pneumoniae, on the surface of the cell.

Example 8

This Example illustrates the expression of truncated and full length PspA by the Mycobacterium tuberculosis strain Bacille Calmette-Guerin (BCG).

BCG was chromosomically modified to incorporate the pspA gene coding for the truncated PspA protein. The 43 kDa truncated PspA protein was expressed from the modified BCG to approximately 15% of total BCG protein. This result was achieved with an expression vector construct carrying the pspA gene segment encoding the 43 kDa region without its 5'-secretion signal. Expression was only about 1% of BCG protein when the PspA or mycobacterial signal sequences were included. In either case, a significant portion of the expressed PspA was excreted into the medium. Expression of the 43 kDa PspA protein in a fusion with the mycobacterial lipoprotein signal sequence resulted in the expression of the recombinant PspA in the membrane fraction of BCG.

This latter result suggested that the fusion of the lipoprotein signal sequence resulted in acylation of the recombinant PspA. Fluorescent activated cell sorting with fluorochrome-conjugated monoclonal antibodies to PspA demonstrated expression of PspA on the surface of these bacteria.

Example 9

This Example illustrates the close homology of portions of the pspA gene among different strains of pneumococci and the potential that such homology may be useful for molecular (DNA) approaches to detect pneumococci in tissues, body fluids and/or secretions and to identify pneumococci grown from patient samples.

Three DNA probes were employed, namely full length pspA, JY4323 (N-terminal HindIII to C-terminal KpnI), the N-terminal half of pspA, JY4306 (N-terminal HindIII to DraI at position 996 (see Figure 3) and most of the proline and

repeat regions, JY4262 (BclI at position 1221 to BstNI at position 1832). Under stringency conditions requiring 95% identity, probes JY4323 and JY4262 reacted with cells of over 200 independent isolates of S.pneumoniae by Southern blot.

When the chromosomal DNA was cut with HindIII, there generally was observed that each of these probes detected two discrete bands whose exact size was dependent on the strain examined. In Rx1 pneumococci, the two bands were 4.0 and 9.1 kb. The 4.0 kb band corresponded to pspA and was altered or absent in pspA mutants. The other band shares some homology with the coding regions for both the N-terminal and C-terminal halves of PspA but is not affected by pspA mutations. The JY4323 and JY4262 probes failed to react with another gram positive bacterium, Streptococcus pyogenes, and a gram negative bacterium, Salmonella typhimurium. The N-terminal probe, JY4306, recognized about one-third of the strains of pneumococci tested.

These results indicate that a sequence included in the proline/repeat region is shared by all strains of pneumococci and apparently not by other bacterial species. Sequences in the N-terminal half of the molecule appear to be more variable.

Example 10

This Example illustrates the detection and determination of the location of epitopes in the α-helical N-terminal region of PspA.

Monoclonal antibodies protective against pneumococcal infection in a mouse model, denoted by an asterisk in Figures 5, 6 and 7, were used to determine the location of epitopes for each antibody in the α-helical N-terminal region of PspA. The sites were mapped to fragments of PspA. The results are illustrated in Figures 5 to 7, with Figure 5 showing the deduced amino acid sequence for the N-terminal region of PspA and the general location of epitopes recognized by monoclonal antibodies, Figure 6 showing antibody reactivity with PspA fragments produced by various pspA gene segments, and Figure 7 showing the mapped location of epitopes in the PspA fragments produced by the different pspA gene segments.

Numbers 138, 193 and 261 in Figure 5 indicate break positions in the PspA fragments used to map the location of epitopes detected by monoclonal antibodies Xi1526, Xi126, XiR35, XiR38, XiR1224, XiR16, Xi64, XiR278, Xi1325 and Xi1323. The asterisk (*) after some of the antibodies denotes those which are able to protect against fatal pneumococcal infection with strain WU2 pneumococci.

In addition, the vertical lines to the right of the Figure indicate those areas predicted to have coiled-coil α-helical structure. The divisions to the left of the Figure indicate the mapped location of the epitopes for each antibody.

## SUMMARY OF DISCLOSURE

In summary of this disclosure, the present invention relates to a truncated PspA molecule capable of eliciting an immunoprotective response and hence containing the protective epitopes of PspA protein. Modifications are possible within the scope of this invention.

## Claims

1. A nucleic acid molecule comprising a nucleotide sequence encoding at least a portion of PspA or having substantial homology to such a nucleotide sequence.

2. The nucleic acid molecule of Claim 1 comprising DNA having a nucleotide sequence included in that set forth in Figure 3, or having substantial homology thereto.

3. The nucleic acid molecule of Claim 1 or Claim 2 coding for at least part of an α-helical coiled coil region of the PspA protein.

4. The nucleic acid molecule of Claim 1 or Claim 2 coding for at least part of a proline rich region of the PspA protein.

5. The nucleic acid molecule of Claim 1 or Claim 2 coding for at least part of a repeat region of the PspA protein.

6. A vector containing the nucleic acid molecule of any one of Claims 1 to 5.

7. A method for obtaining PspA or a protein having substantial homology thereto comprising recovering the PspA or protein having substantial homology thereto from a host organism transformed with the vector of Claim 6.

8. PsPa or a protein having substantial homology thereto obtainable by the method of Claim 7.

9. A vaccine comprising the PspA of Claim 8 and a carrier or diluent.

10. A DNA probe for the detection of pneumococcal DNA in a sample, comprising a conserved sequence of the pspA gene.

11. The DNA probe of Claim 10, wherein said conserved sequence is located in the proline-rich region of the gene, the repeat region of the gene or both.

12. A DNA primer for effecting polymerase chain reactions of pneumococcal DNA comprising a conserved sequence of the pspA gene for assaying a sample for the presence of pneumococci.

13. The DNA primer of Claim 12, wherein said conserved sequence is located in the proline-rich region of the gene, the repeat region of the gene or both.

14. The DNA primer of Claim 13, wherein said conserved sequence comprises a 678 base pair fragment of S. pneumoniae strain Rx1 extending from nucleotide 1312 to 1990 of the pspA sequence of Figure 3.

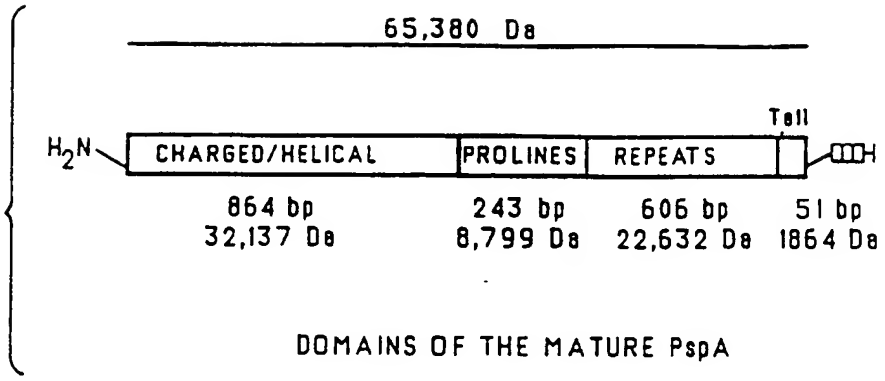


FIG.1.

| | | | | | | | a | b | c | d | e | f | g | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GLU | GLU | ser | pro | val | ala | ser | gln | ser | LYS | ala | GLU | LYS | ASP | 14 |
| | | | | | | | tyr | ASP | ala | ala | LYS | LYS | ASP | 21 |
| | | | | | | | ala | LYS | asn | ala | LYS | LYS | ala | 28 |
| | | | | | | | val | GLU | ASP | ala | gln | LYS | ala | 35 |
| | | | | | | | leu | ASP | ASP | ala | LYS | ala | ala | 42 |
| | | | | | | | gln | LYS | LYS | | | | | 45 |

FIG.2.

FIG.3B.

pspe - sequence -> 1-phase Translation
DNA and derived amino acid 2086 b.p. AAGCTTATGATA.....TCTTTAGGTACC linear

```
  1   /   1                                                  31   /   11
AAG CCT ATG ATA TAG AAA TTT GTA ACA AAA ATG TAA TAT AAA ACA CTT GAC AAA TAT TTA

 61   /   21                                                 91   /   31
CGG AGG AGG CTT ATA CTT AAT ATA AGT ATA GTC TGA AAA TGA CTA TCA GAA AAG AGG TAA

121   /   41                                                151   /   51
ATT TAG ATG AAT AAG AAA AAA ATG ATT TTA ACA AGT CTA GCC AGC GTC GCT ATC TTA GGG
        met asn lys lys lys met ile leu thr ser leu ala ser val ala ile leu gly
181   /   61                                                211   /   71
GCT GGT TTT GTT GCG TCT CAG CCT ACT GTT GTA AGA GCA GAA GAA TCT CCC GTA GCC AGT
ala gly phe val ala ser gln pro thr val val arg ala glu glu ser pro val ala ser
241   /   81                                                271   /   91
CAG TCT AAA GCT GAG AAA GAC TAT GAT GCA GCG AAG AAA GAT GCT AAG AAT GCG AAA AAA
gln ser lys ala glu lys asp tyr asp ala ala lys lys asp ala lys asn ala lys lys
301   /   101                                               331   /   111
GCA GTA GAA GAT GCT CAA AAG GCT TTA GAT GAT GCA AAA GCT GCT CAG AAA AAA TAT GAC
ala val glu asp ala gln lys ala leu asp asp ala lys ala ala gln lys lys tyr asp
361   /   121                                               391   /   131
GAG GAT CAG AAG AAA ACT GAG GAG AAA GCC GCG CTA GAA AAA GCA GCG TCT GAA GAG ATG
glu asp gln lys lys thr glu glu lys ala ala leu glu lys ala ala ser glu glu met
421   /   141                                               451   /   151
GAT AAG GCA GTG GCA GCA GTT CAA CAA GCG TAT CTA GCC TAT CAA CAA GCT ACA CAC AAA
asp lys ala val ala ala val gln gln ala tyr leu ala tyr gln gln ala thr asp lys
481   /   161                                               511   /   171
GCC GCA AAA GAC GCA GCA GAT AAG ATG ATA GAT GAA GCT AAG AAA CGC GAA GAA GAG GCA
ala ala lys asp ala ala asp lys met ile asp glu ala lys lys arg glu glu glu ala
541   /   181                                               571   /   191
AAA ACT AAA TTT AAT ACT GTT CGA GCA ATG GTA GTT CCT GAG CCA GAG CAG TTG GCT GAG
lys thr lys phe asn thr val arg ala met val val pro glu pro glu gln leu ala glu
601   /   201                                               631   /   211
ACT AAG AAA AAA TCA GAA GAA GCT AAA CAA AAA GCA CCA GAA CTT ACT AAA AAA CTA GAA
thr lys lys lys ser glu glu ala lys gln lys ala pro glu leu thr lys lys leu glu
```

FIG.3b.

661 / 221
GAA GCT AAA GCA AAA TTA GAA GAG GCT GAG
glu ala lys ala lys leu glu glu ala glu
721 / 241
GAT GCT GAA GAA GTC GCT CCT CAA GCT AAA
asp ala glu glu val ala pro gln ala lys
781 / 261
CTA GAA CAA GAG CTC AAA GAG ATT GAT GAG
leu glu gln glu leu lys glu ile asp glu
841 / 281
TTC CCT GCT CCT CTT CAA TCT AAA TTG GAT
phe arg ala pro leu gln ser lys leu asp
901 / 301
CAG TTA AGT GAT AAG ATT GAT GAG TTA GAC
glu leu ser asp lys ile asp glu leu asp
961 / 321
AAA GCT GCT GAA GAA AAC AAT AAT GTA GAA
lys ala ala glu glu asn asn asn val glu
1021 / 341
ATT GCT GCT AAA AAA GCT GAA TTA GAA AAA
ile ala ala lys lys ala glu leu glu lys
1081 / 361
GAG CCA GAA AAA CCA GCT CCA GCT CCA GAA
glu pro glu lys pro ala pro ala pro glu
1141 / 381
CCA AAA CCA GCG CCG GCT CCT CAA CCA GCT
pro lys pro ala pro ala pro gln pro ala
1201 / 401
CAA CCA AAA CCA CAA AAA ACA GAT GAT CAA
gln pro lys pro glu lys thr asp asp gln
1261 / 421
GAA GAA GAA TAT AAT GGC TTG ACT CAA CAG
glu glu glu tyr asn arg leu thr gln gln
1321 / 441
GCA CCA CCA ACA GGC TGG AAA CAA GAA AAC
ala pro lys thr gly trp lys gln glu asn

691 / 231
AAA AAA GCT ACT GAA GCC AAA CAA AAA GTG
lys lys ala thr glu ala lys gln lys val
751 / 251
ATC GCT GAA TTG GAA AAT CAA GTT CAT AGA
ile ala glu leu glu asn gln val his arg
811 / 271
TCT GAA TCA GAA CAT TAT GCT AAA GAA GGT
ser glu ser glu asp tyr ala lys glu gly
871 / 291
GCG AAA AAA GCT AAA CTA TCA AAA CTT GAA
ala lys lys ala lys leu ser lys leu glu
931 / 311
GCT GAA ATT GCA AAA CTT GAA GAT CAA CTT
ala glu ile ala lys leu glu asp gln leu
991 / 331
GAC TAC TTT AAA GAA GGT TTA GAG AAA ACT
asp tyr phe lys glu gly leu glu lys thr
1051 / 351
ACT GAA GCT GAC CTT AAG AAA GCA GTT AAT
thr glu ala asp leu lys lys ala val asn
1111 / 371
ACT CCA GCC CCA GAA CCA CCA GCT GAA CAA
thr pro ala pro glu ala pro ala glu gln
1171 / 391
CCC GCA CCA AAA CCA GAG AAG CCA GCT GAA
pro ala pro lys pro glu lys pro ala glu
1231 / 411
CAA CCT GAA GAA GAC TAT GCT CGT AGA TCA
gln ala GLU glu asp tyr ala arg arg ser
1291 / 431
CAA CCG CCA AAA CTT CAA AAA CCA GCT CCT
gln pro pro lys ala glu lys pro ala pro
1351 / 451
GGT ATG TGG TAC TTC TAC AAT ACT GAT GGT
gly met trp tyr phe tyr asn thr asp gly



EP 0 786 521 A1

FIG.3c

| 1381 | / | 461 | | | | | | | | 1411 | / | 471 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TCA | ATG | GCG | ACA | GGA | TGG | CTC | CAA | AAC | AAC | GGT | TCA | TGG | TAC | TAC | CTC | AAC | AGC | AAT | GGT |
| ser | met | ala | thr | gly | trp | leu | gln | asn | asn | gly | ser | trp | tyr | tyr | leu | asn | ser | asn | gly |
| 1441 | / | 481 | | | | | | | | 1471 | / | 491 | | | | | | | |
| GCT | ATG | GCT | ACA | GGT | TGG | CTC | CAA | TAC | AAT | GGT | TCA | TGG | TAT | TAC | CTC | AAC | GCT | AAC | GGC |
| ala | met | ala | thr | gly | trp | leu | gln | tyr | asn | gly | ser | trp | tyr | tyr | leu | asn | ala | asn | gly |
| 1501 | / | 501 | | | | | | | | 1531 | / | 511 | | | | | | | |
| GCT | ATG | GCA | ACA | GGT | TGG | GCT | AAA | GTC | AAC | GGT | TCA | TGG | TAC | TAC | CTC | AAC | GCT | AAT | GGT |
| ala | met | ala | thr | gly | trp | ala | lys | val | asn | gly | ser | trp | tyr | tyr | leu | asn | ala | asn | gly |
| 1561 | / | 521 | | | | | | | | 1591 | / | 531 | | | | | | | |
| GCT | ATG | GCT | ACA | GGT | TGG | CTC | CAA | TAC | AAC | GGT | TCA | TGG | TAT | TAC | CTC | AAC | GCT | AAC | GGC |
| ala | met | ala | thr | gly | trp | leu | gln | tyr | asn | gly | ser | trp | tyr | tyr | leu | asn | ala | asn | gly |
| 1621 | / | 541 | | | | | | | | 1651 | / | 551 | | | | | | | |
| GCT | ATG | GCA | ACA | GGT | TGG | GCT | AAA | GTC | AAC | GGT | TCA | TGG | TAC | TAC | CTC | AAC | GCT | AAT | GGT |
| ala | met | ala | thr | gly | trp | ala | lys | val | asn | gly | ser | trp | tyr | tyr | leu | asn | ala | asn | gly |
| 1681 | / | 561 | | | | | | | | 1711 | / | 571 | | | | | | | |
| GCT | ATG | GCT | ACA | GGT | TGG | CTC | CAA | TAC | AAC | GGT | TCA | TGG | TAC | TAC | CTC | AAC | GCT | AAC | GGT |
| ala | met | ala | thr | gly | trp | leu | gln | tyr | asn | gly | ser | trp | tyr | tyr | leu | asn | ala | asn | gly |
| 1741 | / | 581 | | | | | | | | 1771 | / | 591 | | | | | | | |
| GCT | ATG | GCT | ACA | GGT | TGG | GCT | AAA | GTC | AAC | GGT | TCA | TGG | TAC | TAC | CTC | AAC | GCT | AAT | GGT |
| ala | met | ala | thr | gly | trp | ala | lys | val | asn | gly | ser | trp | tyr | tyr | leu | asn | ala | asn | gly |
| 1801 | / | 601 | | | | | | | | 1831 | / | 611 | | | | | | | |
| GCT | ATG | GCA | ACA | GGT | TGG | GTG | AAA | GAT | GGA | GAT | ACC | TGG | TAC | TAT | CTT | GAA | GCA | TCA | GGT |
| ala | met | ala | thr | gly | trp | val | lys | asp | gly | asp | ser | trp | tyr | tyr | leu | glu | ala | ser | gly |
| 1861 | / | 621 | | | | | | | | 1891 | / | 631 | | | | | | | |
| GCT | ATG | AAA | GCA | AGC | CAA | TGG | TTC | AAA | GTA | TCA | GAT | AAA | TGG | TAC | TAT | GTC | AAT | GGT | TTA |
| ala | met | lys | ala | ser | gln | trp | phe | lys | val | ser | asp | lys | trp | tyr | tyr | val | asn | gly | leu |
| 1921 | / | 641 | | | | | | | | 1951 | / | 651 | | | | | | | |
| GCT | GCC | CTT | GCA | GTC | AAC | ACA | ACT | GTA | GAT | GGC | TAT | AAA | GTC | AAT | GCC | AAT | GGT | GAA | TGG |
| gly | ala | leu | ala | val | asn | thr | thr | val | asp | gly | tyr | lys | val | asn | ala | asn | gly | glu | trp |
| 1981 | / | 661 | | | | | | | | 2011 | / | 671 | | | | | | | |
| GTT | TAA | GCC | GAT | TAA | ATT | AAA | GCA | TGT | TAA | GAA | CAT | TTG | ACA | TTT | TAA | TTT | TGA | AAC | AAA |
| val | OCH | ala | asp | OCH | ile | lys | ala | cys | OCH | glu | his | leu | thr | phe | OCH | phe | OPA | asn | lys |
| 2041 | / | 681 | | | | | | | | 2071 | / | 691 | | | | | | | |
| GAT | AAG | CTT | CGA | TTG | AAT | AGA | TTT | ATG | TTC | GTA | TTC | TTT | AGG | TAC | | | | | |
| asp | lys | val | arg | leu | asn | arg | phe | met | phe | val | phe | phe | tyr | tyr | | | | | |



EP 0 786 521 A1

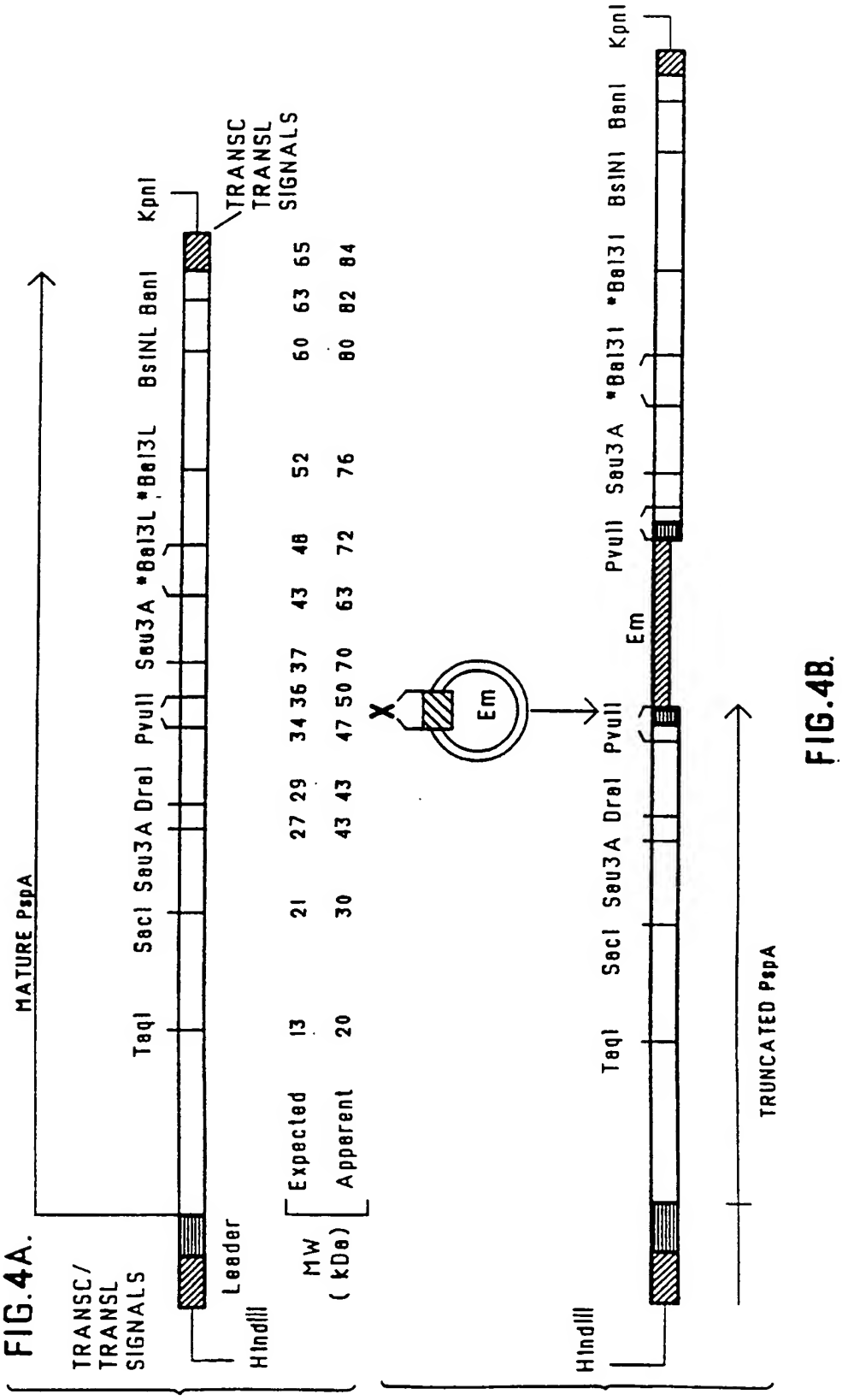
MATURE PspA
TRANSC/
TRANSL
SIGNALS
Leader
HindIII
TaqI
SacI
Sau3A
DraI
PvuII
Sau3A
*Bal3L
*Bal3L
BsINL
BanI
KpnI
TRANSC
TRANSL
SIGNALS
MW
(kDa)
Expected
Apparent
13
20
21
30
27 29
43 43
34 36 37
47 50 70
43
63
48
72
52
76
60
80
63
82
65
84
X
Em
Em
TaqI
SacI
Sau3A
DraI
PvuII
PvuII
Sau3A
*Bal31
*Bal31
BsINI
BanI
KpnI
HindIII
TRUNCATED PspA

FIG.4A.

FIG.4B.

Location of epitopes detected by monoclonal antibodies to PspA

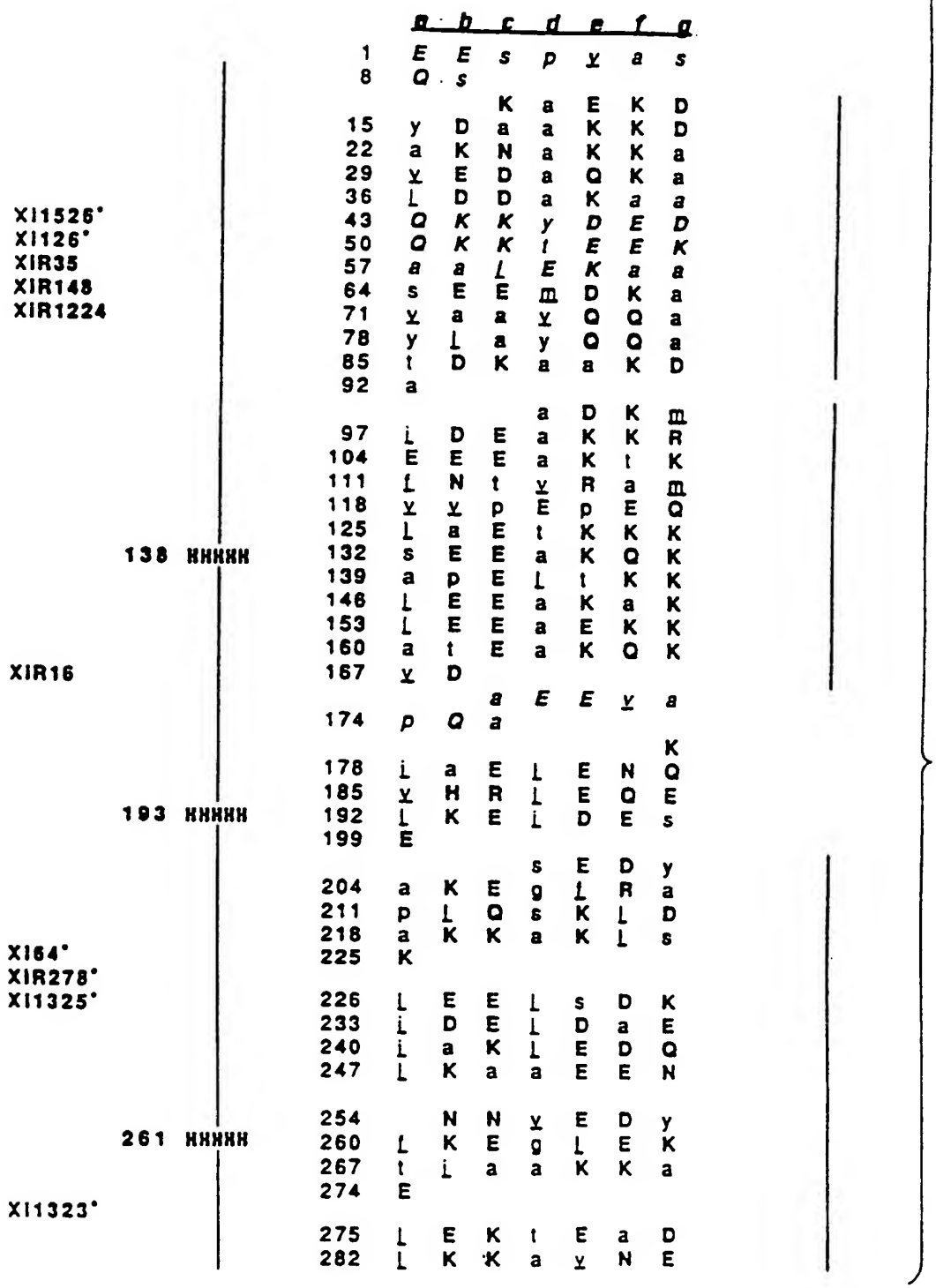


FIG.5.

ANTIBODY REACTIVITY

| | Xi 126 | XiR 1224 | XiR 148 | XiR 1526 | XiR 35 | XiR 16 | XiR 278 | XiR 1325 | XiR 64 | XiR 1323 |
|---|---|---|---|---|---|---|---|---|---|---|
| KSD 1014 | ++ | ++ | ++ | ++ | ++ | ++ | ++ | ++ | ++ | ++ |
| JY 4306 | ++ | ++ | ++ | ++ | ++ | ++ | + | ++ | ++ | ++ |
| JY 4310 | ++ | + | ++ | ++ | ++ | ++ | - | - | - | - |
| JY 4285 | ++ | + | ++ | ++ | ++ | + | - | - | - | - |
| KSD 1500 | - | - | - | - | - | - | - | - | - | - |
| BC 100 | - | - | - | - | - | - | ++ | ++ | ++ | ++ |
| BC 207 | - | - | - | - | - | + | ++ | ++ | ++ | ++ |

FIG.6.

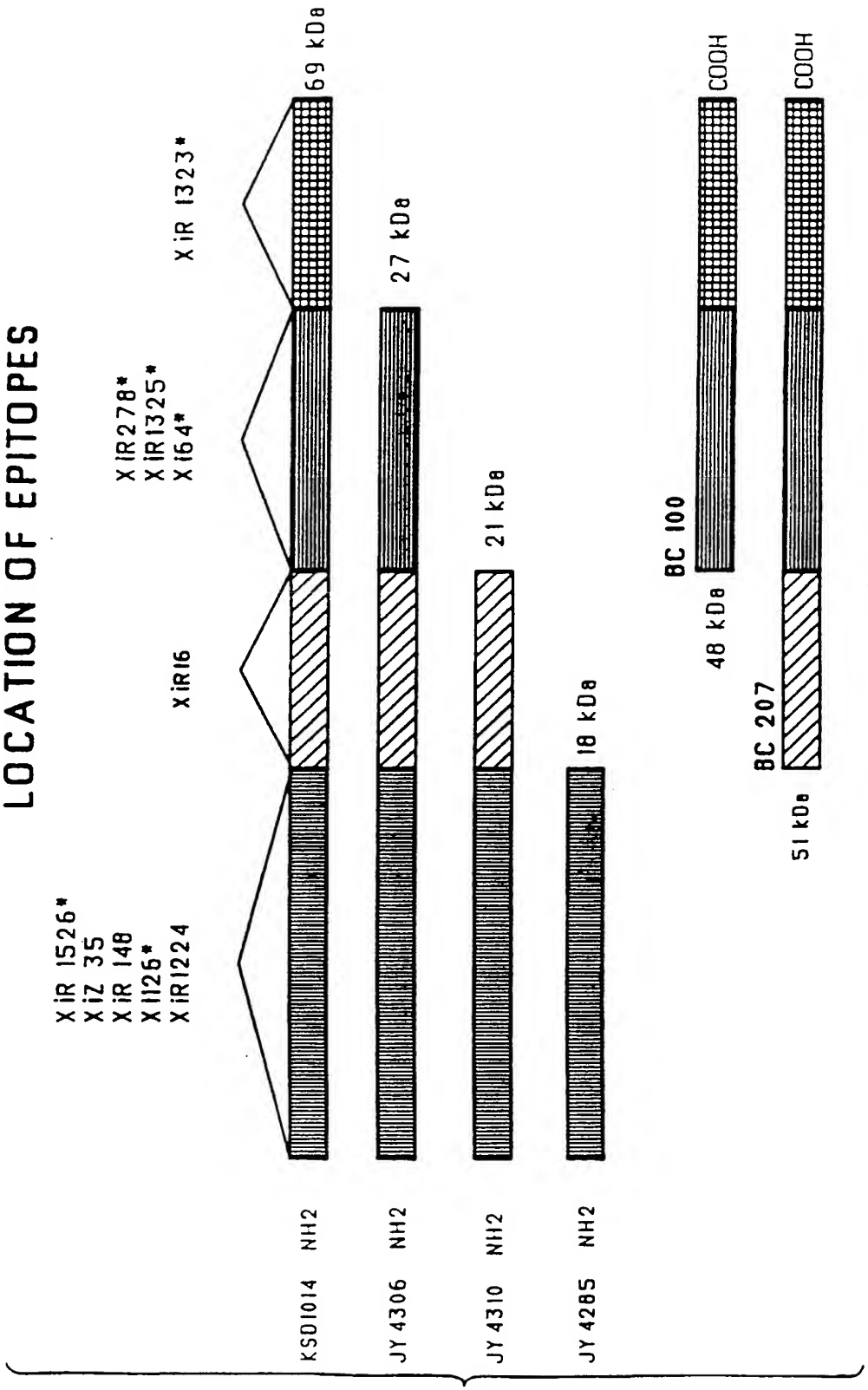
LOCATION OF EPITOPES
X iR 1526*
X iZ 35
X iR 148
X 1126*
X iR1224
X iR16
X iR278*
X iR1325*
X i64*
X iR 1323*
69 kDa
27 kDa
21 kDa
18 kDa
KSD1014 NH2
JY 4306 NH2
JY 4310 NH2
JY 4285 NH2
BC 100
48 kDa
COOH
BC 207
51 kDa
COOH

FIG.7.

European Patent Office

# EUROPEAN SEARCH REPORT

Application Number
EP 96 11 9668

## DOCUMENTS CONSIDERED TO BE RELEVANT

| Category | Citation of document with indication, where appropriate, of relevant passages | Relevant to claim | CLASSIFICATION OF THE APPLICATION (Int.Cl.6) |
|---|---|---|---|
| D,X | INFECTION AND IMMUNITY, vol. 59, no. 1, January 1991, WASHINGTON US, pages 222-228, XP000579731 LARRY S. MCDANIEL ET AL.: "PspA, a surface protein of Streptococcus pneumoniae, is capable of eliciting protection against Pneumococci of more than one capsular type" * abstract * * page 222, right-hand column, paragraph 4 * * page 223, left-hand column, paragraph 4 * * page 224, left-hand column, paragraph 4 - page 225, left-hand column, paragraph 1 * * page 225, right-hand column, paragraph 2 * * page 226, right-hand column, paragraph 3 - paragraph 4 * --- | 1,6-9 | C12N15/31 C12N15/74 C12N15/70 C07K14/315 C12N1/21 A61K39/09 C12Q1/68 |
| D,X | ABSTRACTS OF THE ANNUAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY, May 1990, WASHINGTON US, page 98 XP002031180 J. YOTHER ET AL.: "Generation of truncated forms of pnaumococcal surface protein A (PspA) and DNA sequencing of pspA" * abstract n. D-106 * --- -/-- | 1,4-9 | TECHNICAL FIELDS SEARCHED (Int.Cl.6) C12N C07K A61K |

The present search report has been drawn up for all claims

| Place of search | Date of completion of the search | Examiner |
|---|---|---|
| THE HAGUE | 21 May 1997 | Montero Lopez, B |

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

# EUROPEAN SEARCH REPORT

Application Number
EP 96 11 9668

## DOCUMENTS CONSIDERED TO BE RELEVANT

| Category | Citation of document with indication, where appropriate, of relevant passages | Relevant to claim | CLASSIFICATION OF THE APPLICATION (Int.Cl.6) |
|---|---|---|---|
| P,D, X | JOURNAL OF BACTERIOLOGY, vol. 174, no. 2, January 1992, pages 601-609, XP002031181 JANET YOTHER ET AL.: "Structural properties and evolutionary relationships of PspA, a surface protein of Streptococcus pneumoniae, as revealed by sequence analysis" * page 602, left-hand column, paragraph 1 * * page 602, right-hand column, paragraph 3 - page 605, right-hand column, paragraph 1; figures 2,3 * --- | 1-14 | |
| P,D, X | INFECTION AND IMMUNITY, vol. 59, no. 4, April 1991, WASHINGTON US, pages 1285-1289, XP000579730 DEBORAH T. TALKINGTON ET AL.: "A 43-kilodalton pneumococcal surface protein, PspA: Isolation, protective abilities, and structural analysis of the amino-terminal sequence" * abstract * * page 1285, left-hand column, last paragraph - right-hand column, paragraph 2 * * page 1286, right-hand column, paragraph 4; figure 2 * * page 1287, left-hand column, last paragraph - page 1288, left-hand column, paragraph 1 * * page 1288, right-hand column, paragraph 2 - paragraph 3 * --- -/-- | 1-3,8,9 | TECHNICAL FIELDS SEARCHED (Int.Cl.6) |

The present search report has been drawn up for all claims

| Place of search | Date of completion of the search | Examiner |
|---|---|---|
| THE HAGUE | 21 May 1997 | Montero Lopez, B |

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

# EUROPEAN SEARCH REPORT

Application Number
EP 96 11 9668

## DOCUMENTS CONSIDERED TO BE RELEVANT

| Category | Citation of document with indication, where appropriate, of relevant passages | Relevant to claim | CLASSIFICATION OF THE APPLICATION (Int.Cl.6) |
|---|---|---|---|
| P,X | ABSTRACTS OF THE ANNUAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY, May 1991, WASHINGTON US, page 97 XP002031182 JANET YOTHER ET AL.: "Cloning and expression of the gene encoding pneumococcal surface protein A (PspA) from Streptococcus pneumoniae into Escherichia coli" * abstract n. D-113 * --- | 1,6-9 | |
| P,X | ABSTRACTS OF THE ANNUAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY, May 1991, WASHINGTON US, page 31 XP002031183 LARRY S. MCDANIEL ET AL.: "Localization of protective epitopes on PspA (pneumococcal surface protein A)" * abstract n. B-35 * ----- | 1,6-9 | |
| | | | TECHNICAL FIELDS SEARCHED (Int.Cl.6) |

The present search report has been drawn up for all claims

| Place of search | Date of completion of the search | Examiner |
|---|---|---|
| THE HAGUE | 21 May 1997 | Montero Lopez, B |

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)